MISSING VALUES IN UNREPLICATED ORTHOGONAL DESIGNS

by

R. Dennis Cook
and
Sanford Weisberg

Technical Report #253

Department of Applied Statistics
University of Minneosta
St. Paul

September 15, 1975

SUMMARY

A new non-iterative procedure based on the minimization of the error sums of squares arising from individually selected error contrasts is developed for the estimation of missing values in unreplicated orthogonal designs. It is shown that, under certain conditions, the weighted and unweighted least squares solutions are identical and that the "treatment" and error sums of squares are independent. Application to two-level factorial designs is discussed in detail.

Some key words: Missing values; least squares; error contrasts; factorial designs, 2^{p-q} designs.

1. INTRODUCTION

In an unreplicated orthogonal design requiring n observations, it may be the case that one or more of the observations is missing at random. While a nonorthogonal analysis based on the observed values is always possible, a commonly employed practical expedient is to fill-in the missing value, usually by minimizing the error sum of squares, and proceeding, with minor adjustments, as if the data were fully observed. For latin squares, this procedure was suggested by Allan and Wishart (1930). Yates (1933) discussed this method for randomized blocks (and hence general factorial designs) and presents the fundamental results on the consequences of this procedure. Wilkinson (1958) provides a good summary of this technique in standard cases.

In this paper, we consider the problem of missing observations within a slightly different framework. Any orthogonal design with n observations can always be represented by n orthonormal contrasts. We suppose that k of these contrasts have zero mean and hence we will have at most k degrees of freedom for estimating error variance. If t observations are missing at random, the estimate of the missing value is found by minimizing the sum of squares for the k contrasts estimates with zero mean. In section 2 of this paper, we derive a new non-iterative procedure for estimating the missing values, estimates of the contrasts not nominated for error, the estimate of error variance, and, where appropriate, find moments and distributions of estimators. In this framework, the estimate of the missing values have a simple and interesting form, which we discuss.

Dividing the data into orthonormal contrasts is particularly appropriate for 2^{p-q} designs, where the division into contrasts is common, but no natural error sum of squares appears. The results of Section 2 are applied to this problem in Section 3. A procedure for this problem proposed by Draper and Stoneman (1964) is also discussed in detail.

Finally, in Section 4, we work an example of a 4 x 5 factorial experiment with only part of the interaction to be nominated for error.

2. DEVELOPMENT

Consider an unreplicated orthogonal design requiring n independent observations given by the vector $\underset{\sim}{Y}=(y_1,\ldots,y_n)'$, with error variance $\sigma^2\underset{\sim}{I}$. (If the design includes split plots, the observations are on a sub-plot basis.) Let $\underset{\sim}{X} = (x_{ij})$ be an n x n orthogonal matrix with first row $(1/\sqrt{n}, 1/\sqrt{n},\ldots, 1/\sqrt{n})$, and let $\underset{\sim}{\Gamma} = (\Gamma_1,\ldots,\Gamma_n)'$ be such that $\underset{\sim}{\Gamma} = \underset{\sim}{X}\underset{\sim}{Y}$, so that in the usual analysis of variance framework each Γ_i is an orthonormal contrast (Γ_1 is the overall mean), and $E(\underset{\sim}{\Gamma}) = \underset{\sim}{X}E(\underset{\sim}{Y})$ and $\mathrm{Var}(\underset{\sim}{\Gamma})=\sigma^2\underset{\sim}{X}\underset{\sim}{X}'=\sigma^2 I$. In some problems, such as 2^{p-q} designs, the specification of the $\underset{\sim}{X}$ matrix is natural; in general, the choice of X is not unique. Typically, some of the contrasts, say the last k, are assumed to have zero mean and the sum of squares associated with them, $\sum_{i=n-k+1}^{n} \Gamma_i^2$ becomes the error sum of squares. The remaining n-k contrasts may be of interest individually, in which case Γ_i is the least squares estimate of $E(\Gamma_i)$ with associated sum of squares Γ_i^2. Alternatively, several of the contrasts may span a linear subspace of interest, in which case the sum of squares for these contrasts can be pooled to provide tests for average effects or interactions. Under the assumption of normality the usual computations lead to (fixed-effect) F-tests.

Now suppose that t ($\leq$k) of the planned observations are missing at random. For convenience, assume that the last t are missing, so that $\underset{\sim}{Y}_1 = (y_1,\ldots,y_{n-t})'$ is observed and $\underset{\sim}{Y}_2 = (y_{n-t+1},\ldots,y_n)'$ is missing. A common and useful strategy in analysing orthogonal designs with missing values is to obtain an estimate $\hat{\underset{\sim}{Y}}_2$ of $E(\underset{\sim}{Y}_2)$ and then use the vector $\hat{\underset{\sim}{Y}} = (\underset{\sim}{Y}_1' \; \hat{\underset{\sim}{Y}}_2')$ as if it were the fully observed data vector. This is the approach we consider.

Let $\hat{\underset{\sim}{\Gamma}} = (\hat{\Gamma}_1,\ldots,\hat{\Gamma}_n)'$ be the estimate of $E(\underset{\sim}{\Gamma})$ using $\hat{\underset{\sim}{Y}}$, i.e. $\hat{\underset{\sim}{\Gamma}} = \underset{\sim}{X}\hat{\underset{\sim}{Y}}$. To estimate $\hat{\underset{\sim}{Y}}_2$, let

$$\tilde{\underset{\sim}{Y}} = \begin{pmatrix} \underset{\sim}{Y}_1 \\ \underset{\sim}{0} \end{pmatrix}$$

be the observed data with 0 filled in for the missing values (filling in zero simplifies the computations, but with slight modification of the following, any constants could be filled in), and let $\underset{\sim}{\gamma} = (\underset{\sim}{\gamma}_1,\ldots,\underset{\sim}{\gamma}_n)' = \underset{\sim}{X}\tilde{\underset{\sim}{Y}}$. Partition the $\underset{\sim}{X}$ matrix into

$$\underset{\sim}{X} = \begin{pmatrix} \underset{\sim}{X}_{11} & \underset{\sim}{X}_{12} \\ \underset{\sim}{X}_{21} & \underset{\sim}{X}_{22} \end{pmatrix}$$

such that $\underset{\sim}{X}_{22}$ is the $k \times t$ matrix corresponding to the k contrasts with zero mean and the t missing observations. Also partition $\underset{\sim}{\gamma}$ into $(\underset{\sim}{\gamma}_1', \underset{\sim}{\gamma}_2')$, where $\underset{\sim}{\gamma}_2$ corresponds to the k contrasts with zero means, e.g. $\underset{\sim}{\gamma}_2 = (\gamma_{n-k+1},\ldots,\gamma_n)'$. It follows that

$$E(\underset{\sim}{\gamma}_2) = - \underset{\sim}{X}_{22}\, E(\underset{\sim}{Y}_2) \tag{2.1}$$

$$\mathrm{Var}(\underset{\sim}{\gamma}_2) = \sigma^2 (I - \underset{\sim}{X}_{22}\underset{\sim}{X}_{22}'). \tag{2.2}$$

Equation (2.1) is a linear relation in $E(\gamma_2)$, the known matrix X_{22} and $E(Y_2)$, the quantity to be estimated. The covariance structure given by equation (2.2) suggests that estimates should be obtained by weighted least squares, using the known matrix $(I - X_{22}X_{22}')$ as weights. On the other hand, the traditional approach to filling in missing values is equivalent to ignoring the covariance structure and using unweighted least squares. However, it is easily verified that as long as $I-X_{22}'X_{22}$ is nonsingular the column spaces spanned by X_{22} and $(I-X_{22}X_{22}')X_{22}$ are the same and, hence, the weighted and unweighted least squares solutions will be identical (Watson, 1967). We shall return to the singular case at a later time.

Provided $(X_{22}'X_{22})^{-1}=S^{-1}$, say, exists (which, in practice must be checked) the least squares estimate $\hat{Y}_2$ of $E(Y_2)$ is given by

$$\hat{Y}_2 = -S^{-1} X_{22}'\gamma_2 = -S^{-1}X_{22}'X_{21}Y_1 . \qquad (2.3)$$

Equation (2.3) provides a new non-iterative equation for estimating missing values in any design for any error term. For its use, the matrices X_{22} and X_{21} must be specified. (This is equivalent to specifying the k error contrasts.) This technique will give the same estimates as those obtained by using the iterative technique proposed by Yates (1933), the covariance approach of Bartlett (1937) and then non-iterative algorithm proposed by Rubin (1972).

The mean and variance of $\hat{Y}_2$ are easily found from (2.1), (2.2) and (2.3) to be

$$E(\hat{Y}_2) = -S^{-1} X_{22}' E(\gamma_2) = E(Y_2) \qquad (2.4)$$

$$\mathrm{Var}(\hat{Y}_2) = \sigma^2(S^{-1} - I) . \qquad (2.5)$$

Since $\underset{\sim}{X}$ is orthogonal, it follows that all the diagonal elements of $\underset{\sim}{S}^{-1}$ are greater than 1 (if $\underset{\sim}{S}^{-1}$ exists), and hence $\mathrm{Var}(\hat{\underset{\sim}{Y}}_2)$ is positive definite.

By virtue of equations (2.1) - (2.5), we are in a position to completely study the effects of using $\hat{\underset{\sim}{Y}}_2$ to estimate the missing values. We shall also make use of the fact that $\underset{\sim}{X}\,\underset{\sim}{X}' = \underset{\sim}{I}$, which implies the relationships

$$\begin{aligned} &\underset{\sim}{X}_{11}\underset{\sim}{X}'_{11} + \underset{\sim}{X}_{12}\underset{\sim}{X}'_{12} = \underset{\sim}{I} \qquad && \underset{\sim}{X}_{11}\underset{\sim}{X}'_{21} + \underset{\sim}{X}_{12}\underset{\sim}{X}'_{22} = \underset{\sim}{0} \\ &\underset{\sim}{X}_{21}\underset{\sim}{X}'_{11} + \underset{\sim}{X}_{22}\underset{\sim}{X}'_{12} = \underset{\sim}{0} \qquad && \underset{\sim}{X}_{21}\underset{\sim}{X}'_{21} + \underset{\sim}{X}_{22}\underset{\sim}{X}'_{22} = \underset{\sim}{I}\ . \end{aligned} \tag{2.6}$$

The vector of contrast estimates can be written as

$$\hat{\underset{\sim}{\Gamma}} = \underset{\sim}{X}\begin{pmatrix}\underset{\sim}{Y}_1\\ \hat{\underset{\sim}{Y}}_2\end{pmatrix} = \underset{\sim}{X}\begin{pmatrix}\underset{\sim}{Y}_1\\ -\underset{\sim}{S}^{-1}\ \underset{\sim}{X}_{22}\underset{\sim}{X}_{21}\underset{\sim}{Y}_1\end{pmatrix} = \underset{\sim}{X}\begin{pmatrix}\underset{\sim}{I}\\ -\ \underset{\sim}{S}^{-1}\underset{\sim}{X}'_{22}\underset{\sim}{X}_{12}\end{pmatrix}\underset{\sim}{Y}_1$$

$$= \underset{\sim}{X}\underset{\sim}{M}\underset{\sim}{Y}_1 \text{ (say).} \tag{2.7}$$

Now, from (2.4), it follows that $\hat{\underset{\sim}{\Gamma}}$ is an unbiased estimate of $E(\underset{\sim}{\Gamma})$ and

$$\mathrm{Var}(\hat{\underset{\sim}{\Gamma}}) = \sigma^2\underset{\sim}{X}\underset{\sim}{M}\underset{\sim}{M}'\underset{\sim}{X}'\ . \tag{2.8}$$

Performing the indicated matrix multiplications and using relations (2.6), we find

$$\mathrm{Var}(\hat{\underset{\sim}{\Gamma}}) = \sigma^2\begin{pmatrix}\underset{\sim}{I} + \underset{\sim}{X}_{12}\underset{\sim}{S}^{-1}\underset{\sim}{X}'_{12} & 0\\ 0 & \underset{\sim}{I} - \underset{\sim}{X}_{22}\underset{\sim}{S}^{-1}\underset{\sim}{X}'_{22}\end{pmatrix}\ . \tag{2.9}$$

Several interesting facts should be noted concerning equation (2.9). First, the matrix $\underset{\sim}{I} - \underset{\sim}{X}_{22}\underset{\sim}{S}^{-1}\underset{\sim}{X}_{22}'$ can be easily shown to be idempotent of rank k-t and hence if the original observations were normally distributed, we have the well known result that the residual sum of squares, RSS $= \sum_{i=n-k+1}^{n} \hat{\Gamma}_i^2$ is such that

$$\text{RSS} \sim \sigma^2 \chi^2_{k-t}. \tag{2.10}$$

In the case t = k, if the matrix X_{22} is of full rank then all the Γ_i assumed to have zero mean are exactly zero, and hence we will have RSS = 0. Using k = t has been suggested by Draper and Stoneman (1964) for 2^{p-q} designs and is discussed in Section 3.

Each of the remaining n-k contrasts $\hat{\Gamma}_i$ are unbiased estimates of $E(\Gamma_i)$. To find the variance of $\hat{\Gamma}_i$, let $\underset{\sim}{X}_i$ be the i-th row of $\underset{\sim}{X}$, and then from (2.9)

$$\text{Var}(\hat{\Gamma}_i) = \sigma^2 \underset{\sim}{X}_i' \underset{\sim}{M}\underset{\sim}{M}' \underset{\sim}{X}_i \tag{2.11}$$

Therefore, under normality $\hat{\Gamma}_i^2$ is distributed as $\sigma^2 \underset{\sim}{X}_i'\underset{\sim}{M}\underset{\sim}{M}'\underset{\sim}{X}_i$ times a non-central chi-square with one degree of freedom and non-centrality parameter $\frac{1}{2}\, E(\Gamma_i)^2$. From (2.9) $\hat{\Gamma}_i^2$ and RSS are independent and hence the ratio

$$F = \frac{k-t}{\underset{\sim}{X}_i'\underset{\sim}{M}\underset{\sim}{M}'\underset{\sim}{X}_i} \; \frac{\hat{\Gamma}_i^2}{\text{RSS}} \tag{2.12}$$

is distributed (under the null hypothesis) as central F with 1 and k-t degrees of freedom. Note that the division $\underset{\sim}{X}_i'\underset{\sim}{M}\underset{\sim}{M}'\underset{\sim}{X}_i$ is required to properly scale the statistic.

Since $\underset{\sim}{I} + \underset{\sim}{X}_{12}\underset{\sim}{S}^{-1}\underset{\sim}{X}_{12}'$ is <u>not</u> idempotent, $\Sigma\hat{\Gamma}_i^2$ will not have a chi-square distribution, as was first pointed out by Yates (1933). Furthermore, the n-k tests implied by equation (2.12) may have substantial correlation, requiring the use of a conservative multiple test procedure such as that based on the Bonferroni inequality.

If the number of missing values is t=1, some of the formulas developed this far can be simplified. For convenience assume y_n to be missing and let $D = \sum_{j=n-k+1}^{n} x_{jn}^2$ and $w_i = x_{in}/\sqrt{D}$. The estimate $\hat{y}_n$ of $E(y_n)$ can be shown to be

$$\hat{y}_n = -\sum_{i=n-k+1}^{n} w_i^2 \left(\frac{\gamma_i}{x_{in}}\right) .^* \tag{2.13}$$

Note in (2.13) that Y is a linear combination of the γ_i/x_{in} with weights w_i^2 that depend <u>only</u> on the relative magnitudes of the x_{in}.

If only one Γ_i is in the error term, we see from (2.13) that $\hat{y}_n = -\gamma_i/x_{in}$. Hence, it follows that the estimate $\hat{y}_n$ obtained by nominating k contrasts (with $\chi_{in} \neq 0$) to error simultaneously is simply a linear combination (with weights w_i^2) of the estimates that would be obtained if each of the k contrasts were nominated separately.

*Footnote: This sum is taken over those i for which $x_{in} \neq 0$.

The second moments of the $\hat{\Gamma}_i$ found by substituting $\hat{y}_n$ for this missing y_n are

$$\begin{aligned} \mathrm{Var}(\hat{\Gamma}_i) &= (1+w_i^2)\sigma^2 \\ \mathrm{Cov}(\hat{\Gamma}_i,\hat{\Gamma}_j) &= w_i w_j \sigma^2 \ . \end{aligned} \tag{2.14}$$

Under normality, with t=1, $\hat{\Gamma}_i^2 \sim (1+w_i)\sigma^2\chi^2(1,\frac{1}{2}E(\Gamma_i)^2)$ (for $i \leq n-k$) and hence $(1+w_i)^{-1}$ controls the loss in power for F tests due to the missing value. In some problems this loss may be substantial.

3. DESIGNS WITH ALL FACTORS AT TWO LEVELS

The results given in Section 2, though applicable to any unreplicated orthogonal design, are most readily applied to designs with all factors at 2 levels. We shall write 2^{p-q} to mean a $1/2^q$ fraction of a 2^p design, and shall denote the factor names by the first p capital letters A,B,C,.... Letting $n=2^{p-q}$ be the number of planned data points, the observations $y_1,y_2,\ldots,y_n$ correspond to treatment combinations in the design and are denoted by lower case letters, where, in Yates notation, the presence of a letter corresponds to a factor at its high level, and absence corresponds to the factor at its low level.

In 2^{p-q} designs, 2^{p-q} orthogonal contrasts are usually chosen so that the i-th contrast, Γ_i, corresponds to a treatment effect or interaction and its aliases in the design. For these designs, the contrasts will be orthonormal if we take $x_{1j}=1/\sqrt{n}$ for $j=1,\ldots,n$, and $x_{ij}=\pm 1/\sqrt{n}$, for $i=2,3,\ldots,n$; $j=1,2,\ldots,n$ where the sign of x_{ij} is determined by the "evens versus odds" rule (Cochran and Cox, 1958, Sec. 5.24): $x_{ij}=+1/\sqrt{n}$ if the number of letters in common in the name of the treatment combination y_i and the lowest order name for the factorial effect Γ_j is odd; $x_{ij} = -1/\sqrt{n}$ if the number of letters in common is even. With these definitions $\underset{\sim}{X} = (x_{ij})$ is an orthogonal $n \times n$ matrix.

Suppose that the last k contrasts are nominated to error and consider first the case with t=1 missing value, say y_n. It is easily shown that the estimate of $E(y_n)$ is

$$\hat{Y} = -\frac{\sqrt{n}}{k} \sum_{i=n-k+1}^{n} \pm \gamma_i \qquad (3.1)$$

where the sign of γ_i is the sign of x_{in}. Note that apart from the orthonormalizing constant $\sqrt{n}$, $\hat{Y}$ is the mean of k estimates of $E(y_t)$ obtained by setting each of the k nominated contrasts to zero one at a time. As we shall show below, the estimates from setting one contrast to zero have a particularly simple form. A formula similar to (3.1) is given by Shearer (1973).

The variance of the missing value estimate is given by

$$\text{Var}(\hat{Y}) = \left(\frac{n-k}{k} \right)\sigma^2. \qquad (3.2)$$

The estimate of an effect Γ_i ($i \leq n-k$) is now

$$\hat{\Gamma}_i = \gamma_i + x_{in} \hat{Y} \qquad (3.3)$$

where $\chi_{in} = \pm 1/\sqrt{n}$, and hence

$$\text{Var}(\hat{\Gamma}_i) = (1+1/k)\sigma^2 \qquad (3.4)$$

$$\text{Cov}(\hat{\Gamma}_i, \hat{\Gamma}_j) = \pm \frac{\sigma^2}{2k} \qquad (3.5)$$

$$\rho(\hat{\Gamma}_i, \hat{\Gamma}_j) = \pm \frac{1}{k+1} \qquad (3.6)$$

where in (3.5) and (3.6) the sign of the expression on the R.H.S. is equal to the sign of $x_{in}x_{jn}$. Under normality, for $i \leq n-k$, $\hat{\Gamma}_i^2$ is distributed as $(1+1/k)\ \sigma^2$ times a non central χ^2 with one degree of freedom and non-centrality parameter $\frac{1}{2}(E(\Gamma_i))^2$. Similarly, the residual sum of squares based on k contrasts, $RSS = \sum_{i=n-k+1}^{n} (\gamma_i+x_{in}Y)^2$ is distributed as $\sigma^2\chi^2$ with k-1 degrees of freedom. Since $\hat{\Gamma}_i^2$ and RSS are uncorrelated, their ratio is a multiple of an F statistic.

Listed in Table 1 are the variances and correlations of estimates resulting from the missing value procedure for k = 1,2,...,10. From (3.4)-(3.6) note that these moments depend only on k, and not on n, the specific missing value, or on the nature of the contrasts set to zero. (For comparative purposes, if no missing values are present, the estimates of effects will be uncorrelated with common variance σ^2.) From the table we see that both the correlation and the variance of the estimates decrease as k increases. The case k=1, for example, will give estimates with double the variance of the complete data case and sizable correlation between estimates. As an extreme example in a 2^{10} experiment, with only one missing value, setting k=1 gives estimates with variance equal to the variance obtained from an experiment with one half the observations, in this case wasting 511 data points.

Table 1 here

As a simple, though general and informative, example, consider a complete 2^3 design, with $n = 2^3 = 8$ observations. The $\underset{\sim}{X}$ matrix is given in Table 2 along with the labels given the various factorial effects. Interactions are written last since these are ordinarily the candidates to be nominated for error.

Table 2 here

For concreteness, assume at y_5=ab is missing. Let k=1 and set ABC to zero. The resulting estimates of the missing value is $\hat{y}_5 = -\gamma_8/x_{85}$. The remaining effect estimates are given by $\hat{\Gamma}_i = \gamma_i + x_{i5}\,\hat{y}_5 = \gamma_i + x_{i5}\,(-\gamma_1/x_{85}) = \gamma_i \pm \gamma_8$ where the sign of γ_i is minus if the sign of y_5 is the same in Γ_i and Γ_8 and is + if the sign of y_5 is different in Γ_i and Γ_8. Since the contrasts (without the missing value) Γ_i and Γ_8 are orthonormal, it follows that $\gamma_i \pm \gamma_8$ will <u>always</u> give coefficients of zero to one-half of the observations and ± 2 to the other half. The rule for determining the sign of γ_8 implies that the missing observation will always get coefficient zero, and hence the observations with ± 2 coefficients will make up a <u>complete</u> half replicate of the design originally intended with defining contrast given by the generalized interaction of the effect being estimated (Γ_i) and the effect set to zero, (here, Γ_8). Thus, in the 2^3 design the estimate of the A effect with ABC = 0 is given from the complete one half replicate with $A \cdot ABC = BC$ as defining contrast i.e., $\hat{\Gamma}_2 = (2[a-(1)] + 2[abc-bc])/\sqrt{n}$. Note that the alias of A is $A \cdot BC = ABC$, which is zero by assumption. In a similar fashion, the estimate of B with $ABC = 0$ is found from the complete one-half replicate with defining contrast $B \cdot ABC = AC$; i.e. $\hat{\Gamma}_3 = (2[b-(1)] + 2[abc \cdot bc])/\sqrt{n}$. The estimates of all the effects and the defining contrasts that determine the half replicate used are listed in Table 3. Note that each estimate is made from a different half replicate.

Table 3 here

Draper and Stoneman (1964) suggested the use of the case k=1 in the following procedure: In turn set each of the possible effects to zero, obtaining estimates of the remaining effects in each case. For each choice of effect to be set to zero, examine the half normal plot of the estimated effects and use as estimates the set that gives the "best" plot. There are serious difficulties with this procedure. The effect that is set to zero becomes an alias to all of the estimated effects. If a low order effect, say B, is set to zero, then B becomes aliased with all estimated effects. Since low order effects in a 2^{p-q} design can be easily masked by a significant interaction, a serious bias in the estimates of all other effects may result.

Also, within each plot, the observations will have correlations of $\pm$ .5 (before ordering), and hence even the null shape of the half-normal plot will depend upon the order of the effects, making interpretation of the plot very difficult. Finally, the plots will be correlated, since correlation between estimates of the same effect with different contrasts set to zero is $\pm .5$, while the correlation between different effects with different contrasts set to zero can equal 0, $\pm .25$ or $\pm .5$.

For k = 1,2,3,4, the explicit estimates of the A effect for a 2^3 factorial with y_4 = ab missing are shown in Table 4 for all cases in which the three main effects are estimable. From (3.1), we see that each estimate shown is the mean of the estimates given on

Table 4 here

the first four lines. For example, the estimate with AB,BC both zero is simply the average of the estimate with BC = 0 and the estimate with AB = 0. Alternatively, each of these estimates is a linear combination of the observed simple effects (a-(1), abc-bc, and ac-c) and an estimate of the fourth simple effect (abc + a - b - c). Any of the estimates in Table 4 may be of use under some circumstances, especially in fractional designs with more than three main effects.

If more than one value is missing, equation (2.3) provides a non-iterative method for filling in the missing values for any combination of t and k. Here, one must check to be sure that the matrix $\underset{\sim}{X}_{22}'\underset{\sim}{X}_{22}$ is in fact invertable; in this regard, see the comments of Draper and Stoneman (1964). The nature of the estimates filled in will depend upon the specific observations that are missing, so that general comments analogous to those for the case t=1 cannot be made.

4. AN r x c FACTORIAL

Suppose we have an r x c factorial experiment in which at least one of the factors is quantitative. In the usual analysis, the sums of squares for rows, columsn, and interactions would be divided into linear, quadratic, cubic, etc. components. The higher degree components of the interaction term, say cubic and above, would then be assumed to be zero and be used for an error term in testing and estimation; thus, if a missing value occurs in the design, the usual technique of setting the entire interaction to error is not appropriate for this problem.

For example, consider a 4 x 5 design in which the row factor (A) is qualitative and the column factor (B) is quantitative and equally spaced. The sums of squares for A, B, and A x B can be broken down into one degree of freedom (orthonormal) contrasts as shown in Table 5. Here, the choice of the three orthonormal contrasts for the A space could be made in any meaningful way, while the contrasts for the B space are the usual orthogonal polynomials of the B totals. The interaction contrasts are then determined by taking an outer product of each A contrast with each B contrast, and we can then identify 3 of the contrasts with each of A x B linear, A x B quadratic, A x B cubic and A x B quartic. The sums of squares for A x B cubic and A x B quartic, with 6 degrees of freedom, might be taken as an error term.

Table 5 here

Suppose that one observation, say y_{32} is missing and that only the A x B quartic terms are to be nominated to error. In this case, from Table 5 we see that

$$X_{22} = \begin{bmatrix} 12/\sqrt{840} \\ 0 \\ 0 \end{bmatrix}$$

which is of full column rank (=1) and hence the expectation of the missing value can be estimated from (2.3). If any two missing values occurred, say, y_{14} and y_{32} then we cannot simply use A x B quartic for error because the matrix

$$\begin{bmatrix} -4/\sqrt{840} & 12/\sqrt{840} \\ 4/\sqrt{420} & 0 \\ -4/\sqrt{140} & 0 \end{bmatrix}$$

has rank 1. In this case, A x B cubic and A x B quartic would be needed to estimate the missing values.

REFERENCES

Allen, F.E. and Wishart, J. (1930). "A method of estimating the yield of a missing plot in field experimental work." J. Agric. Sci. 20, 399-406.

Bartlett, M.S. (1937). "Some examples of statistical methods of research in agriculture and applied botany." J.R. Stat. Soc. B., 4, 137-170.

Cochran, W. and Cox, G. (1957). Experimental Designs. Wiley.

Draper, N.R. and Stoneman, D.M. (1964). "Estimating missing values in unreplicated two-level factorial and fractional factorial designs." Biometrics, 20, 443- .

Rubin, D.B. (1972) "A non-iterative algorithm for least squares estimation of missing values in any analysis of variance design." Appl. Statist. 21, 136-141.

Shearer, P.R. (1973). "Missing data in quantitative designs." Appl. Statist., 21, 136-141.

Watson, G.S. (1967). "Linear least squares regression." Ann. Math. Statist., 38, 1679-99.

Wilkinson, G.N. (1958). "Estimation of missing values for the analysis of incomplete data." Biometrics, 14, 257-286.

Yates, F. (1933). "The analysis of replicated experiments when the field results are incomplete." Emp. J. Exp. Agric., 1, 129-142.

Table 1. Variance ($\div\sigma^2$) and correlation between estimated effects when one observation is missing in a 2^{p-q} design.

k	Variance	Correlation
1	2.00	.50
2	1.50	.33
3	1.33	.25
4	1.25	.20
5	1.20	.16
6	1.16	.14
7	1.14	.12
8	1.12	.11
9	1.11	.10
10	1.10	.09

Table 2. $\underset{\sim}{X}$ for a 2^3 design*

factorial effect	y_1=(1)	y_2=a	y_3=b	y_4=c	y_5=ab	y_6=ac	y_7=bc	y_8=abc
Γ_1=mean	+1	+1	+1	+1	+1	+1	+1	+1
Γ_2=A	-1	+1	-1	-1	+1	+1	-1	+1
Γ_3=B	-1	-1	+1	-1	+1	-1	+1	+1
Γ_4=C	-1	-1	-1	+1	-1	+1	+1	+1
Γ_5=AB	+1	-1	-1	+1	+1	-1	-1	+1
Γ_6=AC	+1	-1	+1	-1	-1	+1	-1	+1
Γ_7=BC	+1	+1	-1	-1	-1	-1	+1	+1
Γ_8=ABC	-1	+1	+1	+1	-1	-1	-1	+1

*all entries should be divided by $\sqrt{8}$

Table 3. Estimates* of effects in a 2^3 design with ab missing and Γ_8=ABC=0

Effect	Defining Contrast	Estimate
A	BC	2[a − (1)] + 2[abc − bc]
B	AC	2[b − (1)] + 2[abc − ac]
C	AB	2[bc − c] + 2[ab − a]
AB	C	2[abc − ac − bc + c]
AC	B	2[ac − a − c + (1)]
BC	A	2[bc − b − c + (1)]

*All estimates should be divided by $\sqrt{8}$ for orthonormalizing

Table 4. Explicit estimates of A effect for a 2^3 factorial design with ab missing.

m	Error Contrasts	Estimate
1	ABC	2[a-(1)] + 2[abc-bc]
1	BC	2[abc+a-b-c]
1	AC	2[abc-bc] + 2[ac-c]
1	AB	2[a-(1)] 2[ac-c]
2	ABC,BC	[a-(1)] + [abc-bc] + [abc+a-b-c]
2	ABC,AC	[a-(1)] + 2[abc-bc] + [ac-c]
2	ABC,AB	2[a-(1)] + [abc-bc] + [ac-c]
2	BC,AC	[abc-bc] + [ac-c] + [abc+a-b-c]
2	BC,AB	[a-(1)] [ac-c] + [abc+a-b-c]
2	AC,AB	[a-(1)] + [abc-bc] + 2[ac-c]
3	ABC,BC,AC	(2[a-(1)] + 4[abc-bc] + 2[ac-c] + 2[abc+a-b-c])/3
3	ABC,BC,AB	(4[a-(1)] + 2[abc-bc] + 2[ac-c] + 2[abc+a-b-c])/3
3	ABC,AC,AB	(4[a-(1)] + 4[abc-bc] + 4[ac-c])/3
3	BC,AC,AB	(2[a-(1)] + 2[abc-bc] + 4[ac-c] + 2[abc+a-b-c])/3
4	ABC,BC,AC,AB	2[a-(1)] + 2[abc-bc] + 2[ac-c] + ½[abc+a-b-c]

Observation number

		11	12	13	14	15	21	22	23	24	25	31	32	33	34	35	41	42	43	44	45	divisor
Mean		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	$\sqrt{20}$
A	A_1	1	1	1	1	1	1	1	1	1	1	−3	−3	−3	−3	−3	1	1	1	1	1	$\sqrt{60}$
	A_2	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	0	0	0	0	0	−2	−2	−2	−2	−2	$\sqrt{30}$
	A_3	1	1	1	1	1	−1	−1	−1	−1	−1	0	0	0	0	0	0	0	0	0	0	$\sqrt{10}$
B	Linear B_1	−2	−1	0	1	2	−2	−1	0	1	2	−2	−1	0	1	2	−2	−1	0	1	2	$\sqrt{40}$
	Quad. B_2	+2	−1	−2	−1	2	+2	−1	−2	−1	2	+2	−1	−2	−1	2	+2	−1	−2	−1	2	$\sqrt{56}$
	Cubic B_3	−1	2	0	−2	1	−1	2	0	−2	1	−1	2	0	−2	1	−1	2	0	−2	1	$\sqrt{40}$
	Quartic B_4	+1	−4	6	−4	1	+1	−4	6	−4	1	+1	−4	6	−4	1	+1	−4	6	−4	1	$\sqrt{280}$
A x B Linear	A_1xB_1	−2	−1	0	1	2	−2	−1	0	1	2	6	3	0	−3	−6	−2	−1	0	1	2	$\sqrt{120}$
	A_2xB_1	2	1	0	−1	−2	2	1	0	−1	−2	0	0	0	0	0	4	2	0	−2	−4	$\sqrt{60}$
	A_3xB_1	−2	−1	0	1	2	2	1	0	−1	−2	0	0	0	0	0	0	0	0	0	0	$\sqrt{20}$
A x B Quad.	A_1xB_2	2	−1	−2	−1	2	2	−1	−2	−1	2	−6	3	6	3	−6	2	−1	−2	−1	2	$\sqrt{168}$
	A_2xB_2	−2	1	2	1	−2	−2	1	2	1	−2	0	0	0	0	0	−4	2	4	2	−4	$\sqrt{84}$
	A_3xB_2	2	−1	−2	−1	2	−2	1	2	1	−2	0	0	0	0	0	0	0	0	0	0	$\sqrt{28}$
A x B Cubic	A_1xB_3	−1	2	0	−2	1	−1	2	0	−2	1	3	−6	0	6	−3	−1	2	0	−2	1	$\sqrt{120}$
	A_2xB_3	1	−2	0	2	−1	1	−2	0	2	−1	0	0	0	0	0	2	−4	0	4	−2	$\sqrt{60}$
	A_3xB_3	−1	2	0	−2	1	1	−2	0	2	−1	0	0	0	0	0	0	0	0	0	0	$\sqrt{20}$
A x B Quartic	A_1xB_4	1	−4	6	−4	1	+1	−4	6	−4	1	−3	12	−18	12	−3	1	−4	6	−4	1	$\sqrt{840}$
	A_2xB_4	−1	4	−6	4	−1	−1	4	−6	4	−1	0	0	0	0	0	−2	8	−12	8	−2	$\sqrt{420}$
	A_3xB_4	1	−4	6	−4	1	−1	4	−6	4	−1	0	0	0	0	0	0	0	0	0	0	$\sqrt{140}$

Table 5. A set of 19 orthogonal contrasts to 4 x 5 design with the second factor qualitative. Each table entry must be divided by the divisor given in the last column to orthonormalize.